Assessment of 2020 Census Data Quality Processes

CSAC Discussant: Kunal Talwar

2020 Census: many challenges

- A decade of extensive planning
- New innovations in data collection

- COVID
- Weather and Wildfire disruptions
- Uncertainty on questions and related communication challenges

- Thanks to JASONs for the quick turnaround study!

Planning for adverse cases

- "[…] whether the plan for the decennial census would produce results that improve on standards set in previous decades" is a minimum bar

- Can be clarified further by requiring the plan to do this even in the presence of various adverse events.

- "Constitutionally, the Decennial Census must favor timely good-faith closure (e.g., the certification of state populations for the purpose of apportionment) over unachievable perfection."

- In case of adverse events such as 2020, Bureau should take the time to achieve acceptable standards of quality. Add more buffer in the future?

- Looking forward, the Bureau could formalize and codify the processes and quality standards that it deems necessary.

Not just accuracy

- Aggregate accuracy is an important measure

- Need to ensure no systemic biases

- New innovations can create biases
  - IRS filers exclude very low-income
  - Internet response harder for areas with low broadband penetration

- Other challenges can create biases
  - COVID harder on parents with young kids
  - Response easier for WFH workers

- Aggregate undercount not a sufficient measure of quality

Check for biases

- "[…]the accepted range over previous decades allows for considerable imperfections, as long as these do not knowingly embody *a priori* biases against individual states or statutorily defined classes of individuals."

- As the impact of innovations and adverse events may be disparate across groups, important to measure quality across groups
  - Data used beyond apportionment

- Statutorily-defined classes may not suffice

- Both data quality checks and PES reports should look at quality measures across different groups of people.

JASON recommendations for discussion

- Timeline : The new timeline seems necessary
- Summarize data quality assessments. Minimize use of privacy budget.
- Compare 2019 Test to 2020 Census self-response rates.
- Report time series on response rates
- Collect and analyze data from control system tools
- Use reports generated by CRAVA
- Transparency around data quality ~~anomalies~~ issues
- 2030: explore more AR use
- Tiered messaging on various aspects of the Bureau's efforts